euRathlon 2015 competition
Benchmarking and Scoring
(draft)

26 August 2015

Version 3.2

This document is subject to change, refinement and development. Please visit www.eurathlon.eu for the current version.

Competition date: 17-25 September 2015

Competition venue: Tor del Sale Area, Piombino, Italy. (42.954234 N, 10.599906 E)

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement n° 601205

Content

Tables

Figures

Acronyms

NA	Not Applicable
ND	No Data
NE	Not Evaluated
NS	Not Specified
L1	Land trials, L1: Reconnaissance in urban structure
L2	Land trials, L2: Mobile Manipulation (valve closing)
S1	Sea trials, S1: Navigation and environmental survey
S2	Sea trials, S2: Leak localization and valve closing
A1	Air trials, A1: A1: Aerial Detection and Mapping
A2	Air trials, A2: Aerial Reconnaissance inside a building
L+A	L+A: Sub-Challenge (Land + Air): Survey the building and search for a missing worker
S+A	S+A: Sub-Challenge (Sea + Air): Inspect pipe for leaks and search for a missing worker
L+S	L+S: Sub-Challenge (Land + Sea): Stem the leak
GC	GC: The Grand Challenge (Land+Sea+Air).
PC	Performance Class
AC	Autonomy Class
OPI	Object of Potential Interest
AUV	Autonomous Underwater Vehicle
UAV	Unmanned Aerial Vehicle
UGV	Unmanned Ground Vehicle
USV	Unmanned Surface Vehicle
ROV	Remotely Operated Vehicle
KP	Key Penalty
V	Vehicle video score
W	Weight

1 Introduction

Benchmarking can be used effectively to measure progress with respect to the state of the art. Benchmarking is well established in other disciplines such as medicine where carefully designed experiments, tailored to remove any statistical and external environment bias, are used to assess the benefits of a new technique or a new drug. In the robotics field, it is less established and there are good reasons for this. First, algorithms are embodied into a specific physical robot with its particular physical, sensing and computing constraints. Second, in any real outdoor scenario, environmental constraints cannot be controlled.

The objective of this deliverable is to devise and propose assessment mechanisms to establish the benchmark in robotic systems as well as to provide well-defined scoring scales to mark competitors (teams) for the 2015 euRathlon Competition. The approach is based on metrics and criteria that are applied to robot performance in specific scenarios where all the robotic platforms face identical challenges under the same contextual conditions.

The benchmarking approach proposed for euRathlon 2015 involves **Task Benchmarking** which benchmarks the performance of robots when they perform the tasks and Functionality Benchmarking which benchmarks various performance aspects of each functionality. Each task can implement multiple functionalities and each functionality can be evaluated across multiple tasks and domains (see details in Section 2). Task benchmarking focuses on task achievements, which are decomposed into a set of sub-goals or subtasks that are (as the main element) used to evaluate the performance of teams for the task. The **Functionality benchmarking** will be done based on the data collected when teams perform the tasks in each domain or combined domains.

Another objective of the work-package is to develop a scoring methodology for the euRathlon scenarios that is (1) fair to the competitors and (2) enables the derivation of benchmarks. The scoring approach is also skewed towards higher levels of autonomy to promote progress beyond the state of the art. The benchmarking and scoring approaches are integrated into a single framework in such a way that the benchmarking results for a task are used to derive the task winner.

The rest of this deliverable is as follows. Section 2 presents the overall approach to benchmarking for the 2015 competition. Section 3 presents the scoring mechanism for each task of the 2015 scenarios and the approach to determine the winner of each task as well as the overall winner. Section 4 presents the set of Functionalities involved in each task of the 2015 competition and benchmarks for each of these Functionalities.

2 Approach to Benchmarking

Inspired by and adapted from the benchmarking approach of RoCKIn [1] we propose a system-level benchmark (i.e. Task Benchmark) and module-level (i.e. Functionality Benchmark) for euRathlon 2015. The **Task Benchmark** evaluates the performance of the integrated robot systems while the **Functionality Benchmark** evaluates the performance of a specific module/functionality of the robot systems. Evaluating only the performance of integrated systems does not necessarily inform how the individual modules are contributing the global performance and examine which aspects of the module need to be improved. On the other hand, good performance at module level does not necessarily guarantee that systems integrating a set of well performing individual modules will perform well as an integrated system. Focusing on module-level evaluations alone is also not sufficient to determine which robot system can achieve a specific task. Combining both system-level and module-level benchmarking enables us to perform a deeper analysis and gain useful insight about the performance, advantages and limitations of the whole robot system.

In euRathlon, because of the unstructured nature of the environment and changes in conditions between experiments, the benchmarks will be relatively coarse. And some of the tasks are not easily measurable (the quality of the underwater map for instance).

2.1 Matrix approach to Task and Functionality Benchmarking

As discussed above and in Section 1, in order to perform a specific task which has a set of goals to reach a robot needs to execute a set of functionalities. The Functionality and Task Benchmark can be represented in a matrix form as in Figure 1 below.

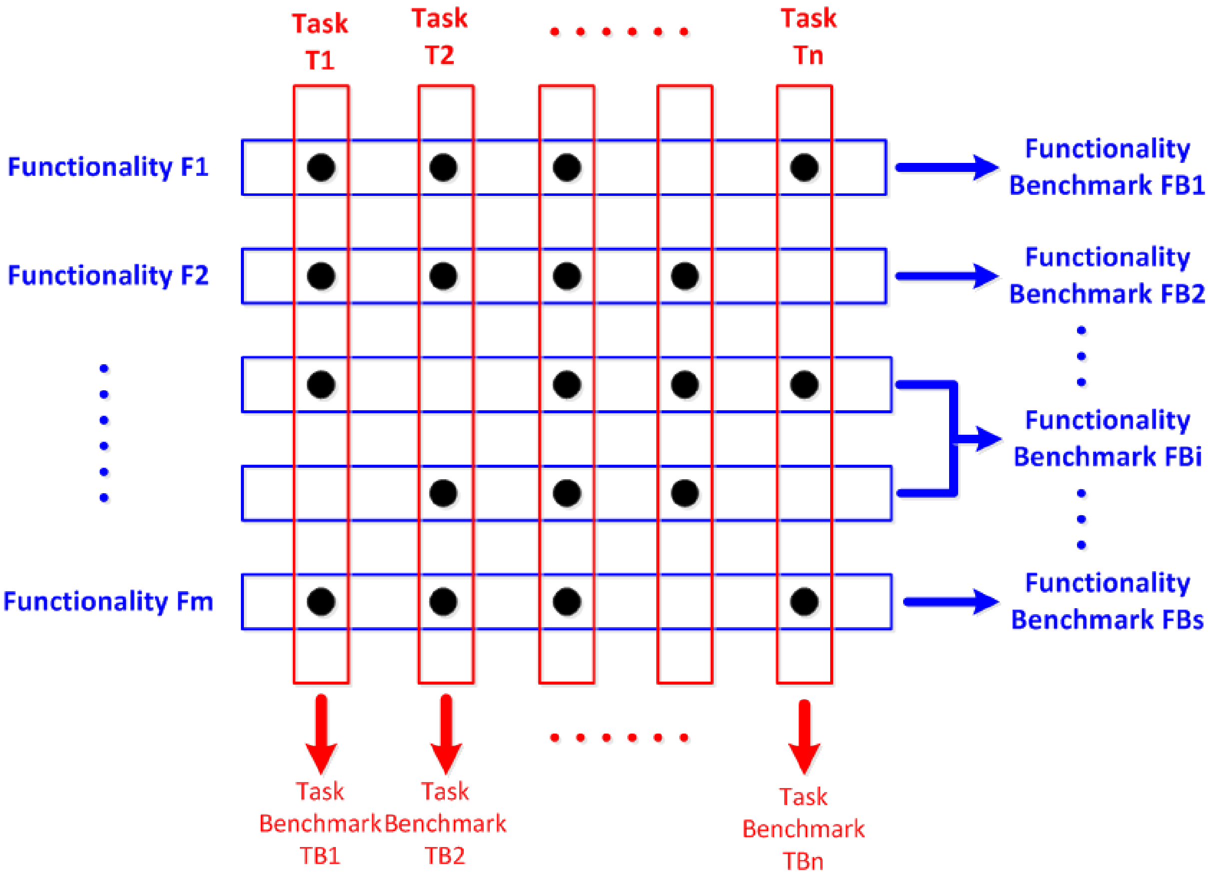

Figure 1. Task (Vertical) and Functionality (Horizontal) Benchmarking illustration[1]

Each task requires the effective implementation of several functionalities to be achieved successfully. Each functionality can be evaluated across different tasks or domains (e.g. Robot Navigation in Land and Sea domain: indoor/outdoor/underwater navigation).

As illustrated in Figure 1, suppose that for the competition we have defined N tasks (T_1, T_2, …, T_n) which correspond to the columns (vertical) and M functionalities (F_1, F_2, …, F_m) which correspond to the rows (horizontal), we will have N Task Benchmarks (TB_1, TB_2, …, TB_n) and S Functionality Benchmarks (here S ≤ M). Because we will benchmark every task there will be the same number of benchmarks as the defined tasks. For some cases it is not quite necessary to evaluate each functionality in a task separately, for instance, a function of Obstacle Avoidance is an essential functionality of a robot but can be considered as part of the Navigation functionality, i.e., one Functionality Benchmark can evaluate more than one functionalities at the same time. This is shown as Functionality Benchmark FB_i in Figure 1. Task (Vertical) and Functionality (Horizontal) Benchmarking illustration above.

The concrete benchmarking for tasks and functionalities are described in detail in Section 3: Task Benchmarking for euRathlon 2015 and Section 4: Functionality Benchmarking for euRathlon 2015.

2.2 Functionality-Task mapping for 2015 scenarios

For the euRathlon 2015 competition, 10 scenarios across 3 domains (Land, Air and Sea) have been defined (more details in [2]). The 10 scenarios are categorised as **Trials** with 2 scenarios in each single domain (as shown in Table 1 below: L1, L2, S1, S2, A1 and A2), **Sub-Challenges** with 3 scenarios in combined two domains (L+A, S+A and L+S) and a **Grand Challenge** (GC) with 3 missions across all three domains. So there are in total 10 tasks corresponding to the 10 scenarios for euRathlon 2015. We have also identified 4 functionalities to be benchmarked as shown in Table 1 below.

Tasks / **Functionalities (/Domain)**	**L1**	**L2**	**S1**	**S2**	**A1**	**A2**	**L+A**	**S+A**	**L+S**	**GC**
2D Mapping (/L+A)	X	-	-	-	X	X	X	X	X	X
Object Recognition (/L+S+A)	X	X	X	X	X	X	X	X	X	X
Obstacle Avoidance (/L+S)	X	X	X	X	-	-	X	X	X	X
Object Manipulation (/L+S)	-	X	-	X	-	-	-	-	X	X

Table 1. Metric representation of the set of tasks and functionalities in euRathlon 2015

The Domain in the Table 1 indicates in which domains the Functionalities are involved.

3 Task Benchmarking for euRathlon 2015

<u>**Task Benchmarking:**</u>

The Task Benchmarking in the euRathlon 2015 competition has a unified framework across all the tasks and is based on the concept of **Performance Class** (**PC**) used as the main element for the ranking of robot performance in a specific task. The Task Benchmarks (framework) consist of following elements: three variables **PC**, **AC** (**Autonomy Class**), **P (Penalties)** and **T** (**Time**) as well as four sets in Table 2. Task Benchmarking class sets definition.

The **Performance Class** (**PC**) measures how well a robot performs in a specific task and is determined by the number of Achievements (or sub-goals) and **Optional Achievements** (**OA**) that the robot achieves during the execution of the task. For example a robot would be in class 3 if it achieves 3 sub-goals out of total 5 sub-goals pre-defined in the task. The higher the class number, the better the performance.

The **Autonomy Class** (**AC**) measures how good a robot performs in terms of Autonomy. In addition to gaining a point for Performance Class, a robot gains an extra autonomy point if a subtask is performed autonomously, or an extra half point if the task is performed semi-autonomously (**Semi-Autonomy: SA**. For example it may happen in the case of manipulation tasks). For example a robot will have a value 4 for **AC** if it performs 4 subtasks autonomously during the execution of the task, or a value 2 for **SA** if semi-autonomously. Similar to PC, the higher the autonomous class number, the better the performance.

set D = **Disqualifying** Behaviours, i.e., things that the robot ***must not do.*** If a disqualifying behaviour is observed the run is terminated with no score and the team is allowed to perform another run with the remaining time allocated to the team slot.
set P = **Penalising** Behaviours, i.e., things that the robot ***should not do***. Some of them are important to the task so that they are defined as Key Penalising behaviours (KP) in the Task benchmarking and will have direct, negative effects on the scoring.
set A = **Achievements** (also called sub-goals), i.e., things that the robot ***should do***.
set OA = **Optional Achievements**, are achievements that are optional objectives, i.e., it is optional for a robot but ***would be better to do***. The set OA could be NULL which indicates that there is no suggested optional sub-goals for a robot to reach. Set **A** and set **OA** are jointly used to determine which Performance Class a robot will be assigned for the specific task. The Achievements of set **A** and **OA** can be either tele-operated or autonomous. If they are performed autonomously, they get Autonomy Class points.
set OPI = Object of Potential Interest, it defines the object the robot ***should detect*** (e.g. a missing worker, a valve etc.) or ***should avoid*** (e.g. some obstacles). It is an important element of the task benchmarks and used in the Achievements for scoring the OPI related operations.

Table 2. Task Benchmarking class sets definition

The **Penalties** (**P**) measures the robustness of a robot. The penalty points are assigned to robots that, in the process of executing the assigned task, make one or more of the errors defined by a task-specific list associated to the Task Benchmark. For example, a robot stops and can't recover by itself and needs direct human manual intervention, one penalty point per intervention.

The **Time** (**T**) records the time a robot takes to accomplish a task. The shorter it is, the better the performance.

Note that for determining the Performance Class of the task an achievement point will ONLY be given if sufficient data is provided for the benchmarking of the functionalities involved in the scenario/task. Required data will be listed in each task benchmark.

Scoring:

The scoring is a part of the benchmarking processes. It is used to rank the teams so as to choose the winning team for task competitions. The task benchmarking for euRathlon 2015 emphasizes task achievements so the scoring will also first consider it. In order to promote autonomy, the scoring will reward teams that perform tasks and sub-tasks autonomously. Therefore, for ranking purposes the following 4-step ranking approach is taken:

Ranking = Performance Score (S) → Autonomy Class(AC) → Penalties (P) → Time (T)

Formula 1 Ranking schema

S = round(PC + 0.5*AC + V - KP) = round (A + OA + 0.5*AC + V – KP)

Formula 2 Performance Score calculation

That is to say, the ranking will first consider the performance score **S**. A team ranks higher with a higher **S**. If two teams have the same **S**, the team with more the autonomy class (**AC**) points ranks higher. Then the penalty points **P** is considered. A team ranks higher with less **P**. If two teams rank the same after **S, AC** and **P**, the team with the shorter time **T** wins.

The performance score **S** considers the performance class points **PC (= A+OA)**. The autonomy class scores **AC** points (one point for Autonomous, half point for semi-autonomous and zero point for tele-operated); the total **AC** is weighed by one half in the final score. The vehicle video scores **V** points: **V** points are awarded to a team when they provide a video describing the vehicle(s); **V** is currently fixed to 2 points. **KP** are the key penalty points: some penalising behaviours (e.g. close wrong valves) are critical to the subtasks so they are labeled as Key Penalty (**KP**) in the Task Benchmarking and will have negative effects in the performance scores.

Value/Points Calculation for Scoring:

In order to score and rank the teams using the approach above for the Task Benchmarking, the value of each of task benchmarking elements need to be calculated. For example, how many points does a team score for its Achievements (**A**), Optional Achievements (**OA**) and Autonomy Class (**AC**), etc.

In order to reflect the nature of different operations with different complexities, each operation (Achievement, Optional Achievement or Penalising behaviour) has an associated weight (**W**) defined in each task for calculating the score of the achievement.

Therefore the general form of the formula for calculating the score for a team for one task is

$$S = \text{round}\left(V + \sum_{i=1}^{n} A_i * W_i + \sum_{j=1}^{m} OA_j * W_j + 0.5 * \sum_{k=1}^{u} AC_k * W_k - \sum_{l=1}^{v} KP_l * W_l \right)$$

Where V is the fixed points for the provided video describing the vehicle(s), n is the number of the Achievements, m is the number of the Optional Achievements, u is the number of achievements (including Achievements and Optional Achievements) done autonomously (one point) or semi-autonomously (half point); v is the number of Key Penalising behaviours. W is their weight defined in the task benchmarks. A_i, OA_j, or AC_k is the base score of the achievement. The base score is 1 in most cases, a real value for some cases such as the map index value and an accumulative value for some cases such as a subtask of "close valves" would have a base score depending on the number valves closed. Examples are given below when describing the base score.

As we can see from the above, we currently do not distinguish **A** from **OA** when computing the score as we think there is no difference between them and this will encourage teams to achieve the **OA** as well while achieving **A**. We keep **OA** here to leave us opportunity for treating them differently when we identify the needs later.

The following rules and/or formulas are used for the calculation of a base score of an achievement:

- Base score for most Achievements

For most achievements/subtasks defined in the benchmarks of each task/scenario, one point is assigned for the base score of each achievement. For example, if a robot achieves two Achievements:

- A1: (W=1) Reaches the area of interest.
- A2: (W=1) Enters the building.

Each of them has a base score of 1 and a weight of 1. So it will get two points: one point for each, therefore the **A** = 2. The **AC** = 2 if they are done autonomously, **AC** = 0 if tele-operated and **AC** = 0.5 *2 = 1 if both subtasks are done semi-autonomously.

Some elements (either Achievements or Penalties) may have points based on the number of occurrences of a specific event. For example,

A5: (W=3) Robots find the missing workers. One point per worker found.

If two missing workers were found, the total points for A5 would be 3 (weight) * 2 = 6

Another example is given below for the valve closing task:

P7: (W=3, KP) The robot closes two or more wrong valves (land or underwater). One penalty per valve.

The notation of "P7: (W=3, KP)" indicates a penalising behaviour has an encoding number P7 with weight 3 for scoring and it is a Key Penalty which will have negative effects on the scoring. E.g. if two wrong valves were closed the score for P7 above would be -3 * 2 = -6.

The base score calculation methods for some special achievements are described as follows.

- Base score for building a map

The quality of the map is used to allocate an index in the range [0,1] derived from the 2D mapping benchmark (see Section 4.1). The final score for it will be **quantized index** multiplied by the weight so that only integer scores are permitted.

The quality factors include

- The covered area
- The number of OPIs (OPI: Object of Potential Interest) marked

For Benchmarking and determining the performance class of the team for this subtask, the quality of the map will be judged based on the ability of the team to use the map to perform the task.

- Base score for detecting OPIs

The number of OPIs detected and the number of false alarms is used to allocate an index in the range [0,1] for the achievement of Detecting/Finding OPIs. A point is allocated for this achievement if the index quality is over a threshold t.

The index is computed based on F-measure as using only Precision and Recall does not provide a single scalar value. The F-measure is defined below in section 4.2. We adopt the traditional balanced F-measure (F_1 measure) that is the harmonic mean of precision and recall. So from the formula (Formula 3) we have

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall}$$

We use the value of F_1 as the value of OPI index. The score for this achievement will be the index value multiplying the weight.

For example, suppose there are in total 10 OPIs on site, a robot detects 8 OPIs but with 2 OPIs misrecognised and 6 false detections. The Precision P = 6 / 8 = 0.75; the Recall R = 6/10 = 0.6, so the F1 = 2*P*R/(P+R) = 2 * 0.75 * 0.6 / (0.75 + 0.6) = 0.67.

The final score for it will be **quantized index** multiplied by the weight so that only integer scores are permitted.

In Section 3.5 to 3.1 below, we will define the benchmarks in detail for each scenario of euRathlon 2015 based on this framework. Some concrete examples to show how the framework works are also given in Sections 3.5 and 3.6. Examples for other scenarios are omitted as they would work in the exact same way.

NOTE: Each time is allocated a fixed time slot for each scenario. If a team is either disqualified during a run or wants to stop a run and perform another trial, they are allowed to do so in the remainder of the time slot allocated to that team. The best run is then used to calculate the score of this team.

3.1 GC: The Grand Challenge (Land+Sea+Air).

Summary of overall objectives (in no specific order):

- From the starting points search for the two workers. Each of them can be on land outside the building, inside the building, on the sea surface near the shore or trapped underwater.
- Inspect the outdoor area and find a safe path to an unobstructed entrance of the building for the UGV. Build a geometric representation of the outdoor area. The path must be shown on the map.
- Enter the building with the UGV and/or UAV and inspect the inside. Build a geometric representation of the building from the inside.
- Find a safe and unobstructed path for the land robot to reach the machine room. The path must be shown on the map.
- The land robot enters in the machine room.

- Reach the underwater pipes area with an AUV (optionally assisted by a USV) surfacing close to the defined waypoints. Follow the plume and find the pipe that is leaking underwater. Build a geometric representation of area.
- Robots report which pipe/s are damaged and leaking.
- When the correct valve underwater is reported by the land robot, look for the corresponding pipe underwater, follow it and find the correct valve underwater to close.
- When the correct valve on land is reported by the underwater robot, look for the correct valve in the machine room to close.
- The land robot and the underwater robot must close the correct valves in a synchronised process (the underwater process must be recorded by the onboard camera of the AUV).
- Localise OPIs that mark obstacles, blocked/unblocked entrances, damages on the wall, the machine room, the missing worker, pipe leaks, valves etc. When an OPI is found, images have to be acquired and provided to the Judges as a proof. The positions of the OPIs with respect to the map(s) built during the sub-challenge must also be provided. No recovery of OPIs is required.
- The aerial robot returns to the landing area, the ground robot returns to the starting point and the AUV surfaces after closing the valve.
- Report data collected to the control station within time. If possible, transmit live position and imagery to the control station.

Timing:

Time limit: 100 minutes. **The data must be provided to the judges within two hours from the end of the team's slot.** (Refer to the Scenario documents for more info)

Benchmarking:

The various sets used in this task are defined as follows.

- Set D (Disqualifying behaviours):
  - D1: The robot damages competition arena.
  - D2: The robot does not conform to safety regulations for the competition.
  - D3: The robot does not have a safety abort function.
  - D4: The robot leaves the flight volumes defined by the organization.
  - D5: The robot enters any of the upper floors of the building.
  - D6: The robot impacts the sensitive dune area.
  - D7: Tele-operation of the USV or AUV except for safety reasons and the manipulation achievement.

  - D8: An AUV surfaces three or more times to get GPS fixes not close to the waypoints.
  - D9: An AUV exits the safety corridor three times.

- Set P (Penalising behaviours)
  - P1: (W=1, KP) The robot needs manual intervention (e.g. it the UAV falls and needs to be manually moved to the take-off area to restart the mission, if the UGV gets stuck and cannot recover by itself, if the AUV gets lost). One penalty per intervention.
  - P2: (W=1) The robot leaves the operating area.
  - P3: (W=1) The robot needs to change batteries or to be refuelled (It doesn't apply to UAV). One penalty for each battery change or refuelling.
  - P4: (W=1, KP) An AUV surfaces at any point not close to the waypoint to get a GPS fix. One penalty point per surface, two times maximum.
  - P5: (W=1, KP) An AUV exits the safety corridor, two times maximum. One penalty point per exit.
  - P6: (W=1, KP) The robot closes one wrong valve (land or underwater).
  - P7: (W=3, KP) The robot closes two or more wrong valves (land or underwater). One penalty per valve.

- Set A (Achievements)
  - A1: (W=3) Robots find the missing workers. One point per worker found.
  - A2: (W=4) Robots build an outdoor map. One map per robot, use best one in each domain for scoring.
  - A3: (W=1) Robots find a safe and unobstructed path to the unblocked entrance of the building.
  - A4: (W=4) Robots build an indoor map of the building. One map per robot, use best one in each domain for scoring.
  - A5: (W=1) A land robot enters the building.
  - A6: (W=1) Robots find a safe and unobstructed path to the machine room.
  - A7: (W=1) A land robot enters in the machine room.
  - A8: (W=1) The robot reaches the underwater piping area from the starting point.
  - A9: (W=4) The robot builds a map of the underwater pipes area.
  - A10: (W=2) The maritime robot surfaces close to the defined navigation waypoint (1 point per each waypoint if it surfaces in a radius of 10 meters, 0.5 in a radius of 10-20 meters, 0 beyond 20 meters).
  - A11: (W=2) The robot finds the underwater leaking pipe and reports it.
  - A12: (W=2) The robot finds the land leaking pipe and reports it.
  - A13: (W=1) The robot follows the pipe.
  - A14: (W=1) The robot inspects the underwater piping assembly where the valve is

mounted.

- A15: (W=2) The robot successfully closes the correct valve in the machine room.
- A16: (W=2) The robot successfully closes the correct valve underwater.
- A17: (W=2) Robots close the correct valves in a synchronized process (land/underwater).
- A18: (W=3) The robots find the OPIs (images must be provided as evidence and the position of the OPI must be also provided and should be shown in the built maps).
- A19: (W=1) The aerial robot(s) returns to the landing area.
- A20: (W=1) The ground robot(s) returns to the starting point.
- A21: (W=1) The AUV surfaces after closing the valve.

- Set OA (Optional Achievements)
  - OA1: (W=1) Robots locate each worker within the first 30 minutes. One point per worker detected.
  - OA2: (W=2) the maps are built on board, and the waypoints of the safe path are also calculated on board (one point of base score for each of three maps).
  - OA3: (W=1) A robot transmits live video images to the control station (one point of base score per domain).
  - OA4: (W=2) Robots cooperate/communicate directly (i.e. not through operators) between domains.
  - OA5: (W=2) The AUV is acoustically supported by the USV.
  - OA6: (W=1) The robots detect all the blocked entrances.
  - OA7: (W=1) A UAV enters the building.
- Set OPI
  - 1 unblocked entrance; 3 or 4 blocked entrances; Y indoor obstacles; Z outdoor obstacles; 2 workers; 1 plume; 1 pipe leak underwater; 2 valves underwater ; 1 damages on the land pipes ; 3 or 4 damages on the building; 1 pipe leak on land ; 1 machine room entrance; 4 valves in the machine room ;

Required Data:

- The built maps (including outdoor, indoor and underwater map with marks of the OPIs)
- Images and positions of OPIs
- The robot communication data
- Timing data

Benchmarked Functionalities:

- 2D Mapping
- Object Recognition
- Obstacle Avoidance

- Object Manipulation

Scoring Examples:

The following are some possible examples of the scoring in a simple faked scenario:

- Set D (Disqualifying behaviours):
  (Omitted)
- Set P (Penalising behaviours)
  (Omitted)
- Set A (Achievements)
  - A1: (W=1) reaches the unobstructed entrance of the building
  - A2: (W=1) enters the building
  - A3: (W=4) maps the area
  - A4: (W=3) identifies all OPIs
  - A5: (W=1) returns to the deployment area
- Set OA (Optional Achievements)
  - OA1: (W=2) builds the map on board.
  - OA2: (W=1) transmits live positions and images/video to the control station.
- Set OPI
  (Omitted)

(Assume that the robots do not perform any of the Disqualifying behaviours as defined in set D).

a) Robot1 takes 30 minutes to finish the whole task and performs all the achievements in Set A and Set OA and performs A1, A2, A4 and OA1 autonomously. No penalties – the robot didn't perform any of Penalising behaviours defined in the set P.

   So the performance score S = round (A + OA + 0.5*AC +V – KP) (refer to Formula 2 Scoring)
   A = A1*1 + A2*1 + round(A3*4) + round(A4*3)+ A5*1.
   Assuming the computed index value is 0.9 for A3 and 0.8 for A4, their respective score is quantized according to their respective weights to obtain an integer number as follows:
   A3: Score = round(A3*4). Here A3*4 = 3.6. round(3.6)=4;
   A4: Score = round(A4*3). Here A4*3 = 2.4. round(2.4)=2;
   then A = 1*1 +1*1 +4 + 2 +1*1 = 10
   OA = OA1*2 +OA2*1 = 1*2 + 1*1 = 3.0
   AC = A1*1 +A2*1 + round(A4*3) + OA1*2 = 1*1 + 1*1 + 2 + 1*2 = 6
   V = 2 (assuming the vehicle video has been provided)
   KP = 0 (no Key Penalising behaviours)
   So the S = round (9.0 + 3.0 + 0.5*6+ 2 – 0) = 17

b) Robot2 does the same but all operations are tele-operated (AC = 0) and the robot needs to be restarted once (one penalty point). It gets S = 14.0. The rank is lower wrt (a) because of a lower score (refer to Formula 1 Ranking schema).

c) Robot3 does the same as Robot1 but it has to be restarted twice, so P=2. Therefore Robot1 ranks higher than Robot3.

d) Robot4 does the same as Robot3 but it takes 25 minutes to finish the task. So Robot4 ranks higher than Robot3.

In summary, the ranking from high to low for the four robots above is
Robot1 > Robot4 > Robot3 > Robot2.

3.2 L+A: Sub-Challenge (Land + Air): Survey the building and search for a missing worker

Summary of overall Objectives (in no specific order):

- From the starting points search for the worker inside and outside the building.
- Find a safe path to an unobstructed entrance of the building for the UGV. Build a geometric representation of the outdoor area. The path must be shown on the map.
- Enter the building with the UGV and/or UAV and inspect the inside. Build a geometric representation of the building from the inside.
- Find a safe and unobstructed path for the land robot to reach the machine room. The path must be shown on the map.
- The land robot enters in the machine room.
- The aerial robot returns to the landing area and the ground robot returns to the starting point.
- Localise OPIs that mark obstacles, blocked/unblocked entrances, damages on the wall, the machine room, the missing worker, etc. When an OPI is found, images have to be acquired and provided to the Judges as a proof. The positions of the OPIs with respect to the map(s) built during the sub-challenge must also be provided. No recovery of OPIs is required.
- Report data collected to the control station within time. If possible, transmit live position and imagery to the control station.

Timing:
Time limit: 45 min. The data must be provided to the judges within one hour from the end of the team's slot. (Refer to the Scenario documents for more info)

Benchmarking:

The various sets used in this task are defined as follows.

- Set D (Disqualifying behaviours):
  - D1: The robot damages competition arena.
  - D2: The robot does not conform to safety regulations for the competition.
  - D3: The robot does not have a safety abort function.
  - D4: The robot leaves the flight volumes defined by the organization.
  - D5: The robot enters any of the upper floors of the building.
  - D6: The robot impacts the sensitive dune area.

- Set P (Penalising behaviours):
  - P1: (W=1, KP) The robot needs manual intervention (e.g. it the UAV falls and needs to be manually moved to the take-off area to restart the mission, if the UGV gets stuck and cannot recover by itself). One penalty per intervention.
  - P2: (W=1) The land robot leaves the operating area.
  - P3: (W=1) The robot needs to change batteries or to be refuelled (It doesn't apply to UAV). One penalty for each battery change or refuelling.

- Set A (Achievements):
  - A1: (W=3) Robots find the missing worker.
  - A2: (W=4) Robots build an outdoor map. One map per robot, use best one in each domain for scoring.
  - A3: (W=1) Robots find a safe and unobstructed path to the unblocked entry of the building.
  - A4: (W=4) Robots build an indoor map of the building. One map per robot, use best one in each domain for scoring.
  - A5: (W=1) A land robot enters the building.
  - A6: (W=1) Robots find a safe and unobstructed path to the machine room.
  - A7: (W=1) A land robot enters the machine room.
  - A8: (W=3) The robots find the OPIs ((images must be provided as evidence and the position of the OPI must be also provided and should be shown in the built maps).
  - A9: (W=1) The aerial robot(s) returns to the landing area.
  - A10: (W=1) The ground robot(s) returns to the starting point.

- Set OA (Optional Achievements)
  - OA1: (W=1) The robots locate the worker within the first 30 minutes.
  - OA2: (W=2) The map is built on board and the waypoints of the safe path are also calculated on board (one point of base score for each of the two maps).

- OA3: (W=1) Robots transmit live positions and images/video to the control station.
- OA4: (W=2) Robots cooperate/communicate directly (i.e. not through operators) between domains.
- OA5: (W=1) The robots detect all the blocked entrances.
- OA6: (W=1) A UAV enters the building.

- Set OPI
  - 1 unblocked entrance; 3 or 4 blocked entrances; Y indoor obstacles; Z outdoor obstacles; W damages on the wall; 1 machine room entrance; 1 worker

Required Data:

- Built maps with OPI marks and positions as well as the path to the machine room.
- OPI images and associated positions.
- Robot communication data
- Timing data

Benchmarked Functionalities:

- 2D Mapping
- Object recognition
- Obstacle Avoidance

3.3 S+A: Sub-Challenge (Sea + Air): Inspect pipe for leaks and search for a missing worker

Summary of overall objectives (in no particular order):

- From the take-off area, inspect the pipes areas. Build a geometric representation of the outdoor areas.
- From the starting point reach the underwater pipes area with a AUV (optionally assisted by a USV) surfacing close to the defined waypoints. Follow the plume and find the pipe that is leaking underwater. Build a geometric representation of area.
- Robots report which pipe/s are damaged and leaking.
- If possible, follow the pipe that is leaking to the piping assembly and inspect it.
- Search for the missing worker
- Localise OPIs that mark damages, pipe leaks, the missing worker, etc. When an OPI is found, images have to be acquired and provided to the Judges as a proof. The positions of the OPIs

with respect to the map(s) built during the sub-challenge must also be provided. No recovery of OPIs is required.

- The aerial robot returns to the landing area and the marine robot surfaces after completing the tasks.
- Report data collected to the control station within time. If possible, transmit live position and imagery to the control station.

Timing:

Time limit: 45 min. The data must be provided to the judges within one hour from the end of the team's slot. (Refer to the Scenario documents for more info)

Benchmarking:

The various sets used in this task are defined as follows.

- Set D (Disqualifying behaviours):
  - D1: The robot damages competition arena.
  - D2: The robot does not conform to safety regulations for the competition.
  - D3: The robot does not have a safety abort function.
  - D4: The robot leaves the flight volumes defined by the organization.
  - D5: The robot enters any of the upper floors of the building.
  - D6: Tele-operation of the USV or AUV except for safety reasons and the manipulation achievement.
  - D7: An AUV surfaces three or more times to get GPS fixes not close to the waypoints.
  - D8: An AUV exits the safety corridor three times.

- Set P (Penalising behaviours)
  - P1: (W=1, KP) The robot needs manual intervention (e.g. it the UAV falls and needs to be manually moved to the take-off area to restart the mission, if the AUV gets lost). One penalty per intervention.
  - P2: (W=1) The robot leaves the operating area.
  - P3 (W=1, KP) An AUV surfaces at any point not close to the waypoint to get a GPS fix. One penalty point per surface, two times maximum.
  - P4 (W=1, KP) An AUV exits the safety corridor, two times maximum. One penalty point per exit.
  - P5: (W=1) The robot needs to change batteries or to be refuelled (It doesn't apply to UAV). One penalty for each battery change or refuelling.

- Set A (Achievements)
  - A1: (W=1) The robot reaches the land piping areas from the take-off area.
  - A2: (W=4) The robot builds an outdoor map of the land pipes area.
  - A3: (W=1) The robot reaches the underwater piping area from the starting point.
  - A4: (W=4) The robot builds a map of the underwater pipes area.
  - A5: (W=2) The maritime robot surfaces close to the defined navigation waypoint (1 point per each waypoint if it surfaces in a radius of 10 meters, 0.5 in a radius of 10-20 meters, 0 beyond 20 meters).
  - A6: (W=2) The robot finds the underwater leaking pipe and reports it.
  - A7: (W=2) The robot finds the land leaking pipe and reports it.
  - A8: (W=3) Robots find the missing work
  - A9: (W=3) The robots find the OPIs (images must be provided as evidence and the position of the OPI must be also provided and should be shown in the built maps).
  - A10: (W=1) The aerial robot(s) returns to the landing area.
  - A11: (W=1) The AUV robot(s) surfaces after completing the tasks.

- Set OA (Optional Achievements)
  - OA1: (W=1) The robots locate the worker within the first 30 minutes.
  - OA2: (W=2) The map is built on board on the robot(s). One per robot.
  - OA3: (W=1) Robots transmit live positions and images/video to the control station. One per robot.
  - OA4: (W=2) The AUV is acoustically supported by the USV.
  - OA5: (W=2) Robots cooperate/communicate directly (i.e. not through operators) between domains.
  - OA6: (W=1) The robot follows the pipe that is leaking to the piping assembly.
  - OA7: (W=1) The robot inspects the piping assembly where the valve is mounted (Evidences required).

- Set OPI
  - X pipe damages on land; 1 pipe leak on land; 1 worker; 1 plume; 1 pipe leak underwater;

Required Data:

- The positions and images of the damaged pipes and the missing worker.
- Maps with the area inspected including the OPIs.
- The robot communication data
- Timing data

Benchmarked Functionalities:

- 2D Mapping (Air robot)
- Object Recognition
- Obstacle Avoidance

3.4 L+S: Sub-Challenge (Land + Sea): Stem the leak

Summary of overall objectives (in no particular order):

- From the building entrance, inspect the land pipes area next to the building. Find the leak on the land pipe. Build a geometric representation of the land pipe area close to the building.
- From the starting point reach the underwater pipes area, follow the plume and find the pipe that is leaking underwater. Build a geometric representation of area.
- Robots must report which pipe/s are leaking so the correct valves can be closed.
- When the correct valve underwater is reported by the land robot, look for the correspondent pipe underwater, follow it and find the correct valve underwater
- When the correct valve on land is reported by the underwater robot, look for the correct valve in the machine room.
- The land robot and the underwater robot must close the correct valves in a synchronised process (the underwater process must be recorded by the onboard camera of the AUV).
- Localise OPIs that mark pipe leaks, valves, etc. When an OPI is found, images have to be acquired and provided to the Judges as a proof. The positions of the OPIs with respect to the map(s) built during the sub-challenge must also be provided. No recovery of OPIs is required.
- The land robot returns to the departure point and the marine robot surfaces after closing the valve.
- Report data collected to the control station within time. If possible, transmit live position and imagery to the control station.

Timing:

Time limit: 60 min. The data must be provided to the judges within one hour from the end of the team's slot. (Refer to the Scenario documents for more info)

Benchmarking:

The various sets used in this task are defined as follows.

- Set D (Disqualifying behaviours):
  - D1: The robot damages competition arena.
  - D2: The robot does not conform to safety regulations for the competition.
  - D3: The robot does not have a safety abort function.
  - D4: The robot impacts the sensitive dune area.

- D5: The robot enters any of the upper floors of the building.
- D6: Tele-operation of the USV or AUV except for safety reasons and the manipulation achievement.
- D7: An AUV surfaces three or more times to get GPS fixes not close to the waypoints.
- D8: An AUV exits the safety corridor three times.

- Set P (Penalising behaviours)
  - P1: (W=1, KP) The robot needs manual intervention (e.g. if the UGV gets stuck and cannot recover by itself, if the AUV gets lost). One penalty per intervention.
  - P2: (W=1) A robot leaves the operating area.
  - P3: (W=1, KP) The robot closes one wrong valve (land or underwater).
  - P4: (W=3, KP) The robot closes two or more wrong valves (land or underwater). One penalty per valve.
  - P5 (W=1, KP) An AUV surfaces at any point not close to the waypoint to get a GPS fix (one penalty point per surface, two times maximum).
  - P6 (W=1, KP) An AUV exits the safety corridor (one penalty point per exit, two times maximum).
  - P7: (W=1) The robot needs to change batteries or to be refuelled. One penalty for each battery change or refuelling.

- Set A (Achievements)
  - A1: (W=1) The robot reaches the land piping area from the starting point.
  - A2: (W=4) The robot builds an outdoor map of the land pipes area close to the building.
  - A3: (W=1) The robot reaches the underwater piping area from the starting point.
  - A4: (W=4) The robot builds a map of underwater pipes area.
  - A5: (W=2) The maritime robot surfaces close to the defined navigation waypoint (1 point per each waypoint if it surfaces in a radius of 10 meters, 0.5 in a radius of 10-20 meters, 0 beyond 20 meters).
  - A6: (W=2) The robot finds the underwater leaking pipe and reports it.
  - A7: (W=2) The robot finds the land leaking pipe and reports it.
  - A8: (W=1) The robot enters the building after inspecting the land pipes.
  - A9: (W=1) The robot enters the machine room.
  - A10: (W=2) The robot successfully closes the correct valve in the machine room.
  - A11: (W=2) The robot successfully closes the correct valve underwater.
  - A12: (W=2) Robots close the correct valves in a synchronized process (land/underwater).

- A13: (W=3) The robots find the OPIs (images must be provided as evidence and the position of the OPI must be also provided and should be shown in the built maps).
- A14: (W=1) The land robot(s) returns the starting point.
- A15: (W=1) The AUV robot(s) surfaces after closing the valve.

- Set OA (Optional Achievements)
    - OA1: (W=2) The map is built on board on the robot(s). One per robot.
    - OA2: (W=1) The robots transmit live positions and images/video to the control station. One per robot.
    - OA3: (W=2) The AUV is acoustically supported by a USV
    - OA4: (W=2) Robots cooperate/communicate directly (i.e. not through operators) between domains.
    - OA5: (W=1) The robot follows the pipe that is leaking to the piping assembly.
    - OA6: (W=1) The robot inspects the piping assembly where the valve is mounted (Evidences required)

- Set OPI
    - 1 pipe leak on land; 1 machine room entrance; Y indoor obstacles; Z outdoor obstacles; 1 valve machine room; 1 plume; 1 pipe leak underwater; 1 valve underwater

Required Data:

- Images and position of OPIs (only the relevant ones will be scored)
- Maps with the OPIs correctly positioned
- The robot communication data
- Timing data

Benchmarked Functionalities:

- 2D Mapping (Land robot)
- Object Recognition
- Obstacle Avoidance
- Object Manipulation

3.5 L1: Reconnaissance in urban structure

Summary of overall objectives (in no particular order):

- From the starting point find and follow a safe path to an unobstructed entrance of the building. Build a geometric representation of the outdoor area (from the starting point to the building).
- Enter the building and inspect the inside. Build a geometric representation of the building from the inside.
- Find a safe and unobstructed path to reach the machine room
- Enter the machine room.
- Return to the deployment area.
- Localise OPIs that mark obstacles, blocked/unblocked entrances and the machine room entrance. When an OPI is found, images have to be acquired and provided to the Judges as a proof. The positions of the OPIs with respect to the map built during the trial must also be provided. No recovery of OPIs is required. The safe path followed by the UGV to reach the unblocked entrance and to reach the machine room should also be shown in the outdoor and indoor maps respectively.
- Report data collected to the control station within the time. If possible, transmit live position and imagery to the control station.

Timing:

Time limit: 45 min. The data must be provided to the judges within one hour from the end of the team's slot. (Refer to the Scenario documents for more info)

Benchmarking:

We first define the various sets used in this task.

- Set D (Disqualifying behaviours):
  - D1: The robot damages competition arena.
  - D2: The robot does not conform to safety regulations for the competition.
  - D3: The robot does not have a safety abort function.
  - D4: The robot impacts the sensitive dune area.
  - D5: The robot enters any of the upper floors of the building.

- Set P (Penalising behaviours)
  - P1: (W=1, KP) The robot needs manual intervention (e.g. if the UGV gets stuck and cannot recover by itself). (One penalty per intervention).

  - P2: (W=1) The robot leaves the operating area.
  - P3: (W=1) The robot needs to change batteries or to be refueled. One penalty for each battery change or refueling.

- Set A (Achievements)
  - A1: (W=1) The robot reaches the unobstructed entrance of the building
  - A2: (W=4) The robot builds an outdoor map (from the starting point to the building).
  - A3: (W=4) The robot builds an indoor map of the building.
  - A4: (W=1) The robot finds a safe unobstructed path to the machine room
  - A5: (W=1) The robot enters the machine room
  - A6: (W=3) The robot finds the OPIs (images must be provided as evidence and the position of the OPI must be also provided and should be shown in the built maps).
  - A7: (W=1) The robot returns the deployment area.

- Set OA (Optional Achievements)
  - OA1: (W=2) The map is built on board on the robot(s).
  - OA2: (W=1) The robot transmits live positions and images/video to the control station.
- Set OPI
  - 1 unblocked entrance; 3 or 4 blocked entrances; 1 machine room entrance; Y indoor obstacles; Z outdoor obstacles

Then the scoring will be done based on the framework described in the beginning of Section 3. The actual scoring sheets and guidelines for the judges will be designed separately.

Required Data:
- The area map (e.g. KML format with ID and Tags for every element of the map including objects found, obstacles, images/video data evidence, please refer to 5. Benchmarking Data Formats).
- Timing data

Benchmarked Functionalities (as defined in Section 4 Functionality Benchmarking for euRathlon 2015**):**
- 2D Mapping
- Object Recognition
- Obstacle Avoidance

3.6 L2: Mobile Manipulation (valve closing)

Summary of overall objectives (in no particular order):

- From the entrance of the building, reach the machine room where the valves and the canister are located.
- Close the valves to stem the leaks.
- Pick up the canister from the ground and drop it into the barrel. If possible, close the barrel.
- Return to the deployment area.
- Localise OPIs that mark obstacles, the machine room entrance, the valves, the canister and the barrel. When an OPI is found, images have to be acquired and provided to the Judges as a proof. No recovery of OPIs is required. Report data collected to the control station within the time specified in the timing section below. If possible, transmit live position and imagery to the control station.
- Report data collected to the control station within time. If possible, transmit live position and imagery to the control station.

Timing:

Time limit: 45 min. The data must be provided to the judges within one hour from the end of the team's slot. (Refer to the Scenario documents for more info)

Benchmarking:

The various sets used in this task are defined as follows.

- Set D (Disqualifying behaviours):
  - D1: The robot damages competition arena.
  - D2: The robot does not conform to safety regulations for the competition.
  - D3: The robot does not have a safety abort function.
  - D4: The robot enters any of the upper floors of the building.

- Set P (Penalising behaviours)
  - P1: (W=1, KP) The robot needs manual intervention (e.g. if the UGV gets stuck and cannot recover by itself). (one penalty per intervention).
  - P2: (W=1) The robot leaves the operating area.
  - P3: (W=1) The robot needs to change batteries or to be refueled. One penalty for each battery change or refueling.

- Set A (Achievements)
  - A1: (W=1) The robot reaches the machine room.
  - A2: (W=2) The robot closes each valve (one point for closing each value)
  - A3: (W=1) The robot picks up the canister.
  - A4: (W=1) The robot drops the canister into the barrel for contaminated materials.
  - A5: (W=3) The robot finds the OPIs (images must be provided as evidence).
  - A6: (W=1) The robot returns the deployment area.

- Set OA (Optional Achievements)
  - OA1: (W=2) The robot closes the barrel.
  - OA2: (W=1) The robot transmits live position and imagery/video to the control station.
- Set OPI
  - 1 machine room entrance; 4 valves ; 1 canister; 1 barrel ; X indoor obstacles

Required Data:

- OPI positions and images/video data
- Timing data

Benchmarked Functionalities:

- Object Recognition
- Obstacle Avoidance
- Object Manipulation

3.7 S1: Navigation and environmental survey

Summary of overall objectives:

- The AUV passes through the gate
- The AUV reaches the area of interest
- Inspect the area and build a geometric representation of the area and its environment.
- Localise OPIs that mark the gate and damages to the seabed. When an OPI is found, images have to be acquired and provided to the Judges as a proof. The positions of the OPIs with respect to the map built during the trial must also be provided. No recovery of OPIs is required.
- Return to the departure point.
- Report data collected to the control station. If possible, transmit live position and imagery/video to the control station.

Timing:
Time limit: 40 min. The data must be provided to the judges within one hour from the end of the team's slot. (Refer to the Scenario documents for more info)

Benchmarking:
The various sets used in this task are defined as follows.

- Set D (Disqualifying behaviours):
  - D1: The robot damages competition arena.
  - D2: The robot does not conform to safety regulations for the competition.
  - D3: The robot does not have a safety abort function.
  - D4: Tele-operation of the USV or AUV except for safety reasons.

- Set P (Penalising behaviours)
  - P1: (W=1) The robot leaves the operating area.
  - P2: (W=1, KP) The robot needs direct human intervention (e.g. if the AUV gets lost). One penalty point per intervention.
  - P3: (W=1) The robot needs to change batteries. One penalty for each battery change.

- Set A (Achievements)
  - A1: (W=1) The AUV passes through the gate.
  - A2: (W=1) The AUV reaches the area of interest.
  - A3: (W=4) The robot builds a map of the area and its environment.
  - A4: (W=3) The robot finds the OPIs (images must be provided as evidence and the position of the OPI must be also provided and should be shown in the built maps).
  - A5: (W=1) The robot returns to the departure point.

- Set OA (Optional Achievements)
  - OA1: (W=1) The robot transmits live positions and images/video to the control station.
  - OA2: (W=2) The AUV is acoustically supported by a USV.
- Set OPI
  - 1 gate; 5 damages on the seabed;

Required Data:
- The area and structure map with OPI marks
- The OPI positions and images/video data
- Timing data

Benchmarked Functionalities:

- Object Recognition
- Object Avoidance

3.8 S2: Leak localization and valve closing

Summary of overall objectives:

- The AUV passes through the gate
- The AUV reaches the pipes area.
- Follow the plume and find the pipe that is leaking.
- Once the piping assembly structure is reached, find the valve.
- Close the valve (the onboard camera must record the operation).
- Surface after closing the valve.
- Build a geometric representation of the area and its environment and localise OPIs that mark the gate, plume, the pipe leak and the valve. When an OPI is found, images have to be acquired and provided to the Judges as a proof. The positions of the OPIs with respect to the map built during the trial must also be provided. No recovery of OPIs is required.
- Report data collected to the control station within time. If possible, transmit live position and imagery to the control station.

Timing:

Time limit: 50 min. The data must be provided to the judges within one hour from the end of the team's slot. (Refer to the Scenario documents for more info)

Benchmarking:

The various sets used in this task are defined as follows.

- Set D (Disqualifying behaviours):
  - D1: The robot damages competition arena.
  - D2: The robot does not conform to safety regulations for the competition.
  - D3: The robot does not have a safety abort function.
  - D4: Tele-operation of the USV or AUV except for safety reasons and for the manipulation task.

- Set P (Penalising behaviours)
  - P1: (W=1) The robot leaves the operating area.
  - P2: (W=1, KP) The robot needs direct human intervention (e.g. if the AUV gets lost). One penalty point per intervention.
  - P3: (W=1) The robot needs to change batteries. One penalty for each battery change.

- Set A (Achievements)
  - A1: (W=1) the AUV passes through the gate.
  - A2: (W=1) The AUV reaches the piping area from the starting point.
  - A3: (W=4) The robot builds a map of the area and its environment.
  - A4: (W=3) The robot finds the OPIs (images must be provided as evidence and the position of the OPI must be also provided and should be shown in the built maps). In the case of the pipe, evidence of pipe following must be shown.
  - A5: (W=4) The robot closes the valve. The onboard camera of the vehicle must also record the operation.
  - A6: (W=1) The AUV surfaces after closing the valve.

- Set OA (Optional Achievements)
  - OA1: (W=2) Imaging of the piping assembly structure.
  - OA2: (W=1) The robot transmits live positions and images/video to the control station.
  - OA3: (W=2) The AUV is acoustically supported by a USV

- Set OPI
  - 1 gate; 1 plume (constituted of 5 buoys) ; 1 pipe leak; 1 valve

Required Data:

- The area and structure map with OPI marks
- Position and images/video data of the pipe, valve and piping assembly structure
- Timing data

Benchmarked Functionalities:

- Object Recognition
- Object Avoidance
- Object Manipulation

3.9 A1: Aerial Detection and Mapping

Summary of overall objectives (in no particular order):

- From the initial point, inspect the outside of the building and its surroundings for building a map of the outdoor area.
- Find a safe and unobstructed path from the departure point of the land robot to an unobstructed entrance of the building. The path must be shown on the map.
- Return to the landing area.

- Localise OPIs that mark obstacles and unblocked/blocked entrances. When an OPI is found, images have to be acquired and provided to the Judges as a proof. The positions of the OPIs with respect to the map built during the trial must also be provided. No recovery of OPIs is required.
- Report data collected to the control station within the time specified in the timing section below. If possible, transmit live position and imagery to the control station.

Timing:

Time limit: 30 min. The data must be provided to the judges within one hour from the end of the team's slot. (Refer to the Scenario documents for more info)

Benchmarking:

The various sets used in this task are defined as follows.

- Set D (Disqualifying behaviours):
  - D1: The robot damages competition arena.
  - D2: The robot does not conform to safety regulations for the competition.
  - D3: The robot does not have a safety abort function.
  - D4: The robot leaves the flight volumes defined by the organization.

- Set P (Penalising behaviours)
  - P1: (W=1, KP) The robot needs manual intervention (e.g. it falls and needs to be manually moved to the take-off area to restart the mission). One penalty per intervention).
  - P2: (W=1) The robot needs to change batteries or to be refuelled more than once (one penalty for each extra battery change or refuelling).
- Set A (Achievements)
  - A1: (W=4) The robot creates an outdoor map of the area.
  - A2: (W=3) The robot finds the OPIs (images must be provided as evidence and the position of the OPI must be also provided and should be shown in the built maps)
  - A3: (W=2) The robot finds an unobstructed path from the UGV deployment area to the unblocked entrance of the building (providing the coordinates of the waypoints)
  - A4: (W=1) The robot returns to the landing area.

- Set OA (Optional Achievements)
  - OA1: (W=1) The robot detects the blocked entrances (in addition to detecting the unblocked entrance). One point per each blocked entrance.
  - OA2: (W=1) The robot finds the shortest safe path to the entrance of the building.
  - OA3: (W=2) The map is built on board the robot(s), and the waypoints of the safe path are also calculated on board.

  - OA4: (W=1) The robot transmits live positions and images/video to the control station.
- Set OPI
  - 1 unblocked entrance; 3 or 4 blocked entrances; Y outdoor obstacles

Required Data:

- Built area map with OPI marks and positions as well as the waypoints of the estimated path.
- OPI images and associated positions.
- Coordinates of the waypoints that define the unobstructed path (at least one waypoint per 2 meters).
- Timing data

Benchmarked Functionalities:

- 2D Mapping
- Object recognition

3.10 A2: Aerial Reconnaissance inside a building

Summary of overall objectives (in no particular order):

- From the initial point enter the building
- Inspect the inside of the building and build a map of the indoor area.
- Find a safe and unobstructed path for a land robot to reach the machine room. (The path must be shown on the map)
- Enter the machine room.
- Return to the initial point.
- Localise OPIs that mark obstacles, damages and blocked paths that would limit access for a ground robot. When an OPI is found, images have to be acquired and provided to the Judges as a proof. The positions of the OPIs with respect to the map built during the trial must also be provided. No recovery of OPIs is required.
- Report data collected to the control station within the time specified in the timing section below. If possible, transmit live position and imagery to the control station.

Timing:

Time limit: 40 min. The data must be provided to the judges within one hour from the end of the team's slot. (Refer to the Scenario documents for more info)

Benchmarking:

The various sets used in this task are defined as follows.

- Set D (Disqualifying behaviours):

- D1: The robot damages competition arena.
- D2: The robot does not conform to safety regulations for the competition.
- D3: The robot does not have a safety abort function.
- D4: The robot leaves the flight volumes defined by the organization.
- D5: The robot enters any of the upper floors of the building.

- Set P (Penalising behaviours)
  - P1: (W=1, KP) The robot needs manual intervention (e.g. it falls and needs to be manually moved to the take-off area to restart the mission). One penalty per intervention.
  - P2: (W=1) The robot needs to change batteries or to be refuelled more than once (one penalty for each extra battery change or refuelling).
- Set A (Achievements)
  - A1: (W=1) The robot enters the building.
  - A2: (W=4) The robot creates an indoor map of the building.
  - A3: (W=2) The robot finds an unobstructed path to the machine room.
  - A4: (W=1) The robot enters the machine room.
  - A5: (W=1) The robot returns to the initial point.
  - A6: (W=3) The robot finds the OPIs (images must be provided as evidence and the position of the OPI must be also provided and should be shown in the built maps).

- Set OA (Optional Achievements)
  - OA1: (W=1) The robot detects all the blocked paths.
  - OA2: (W=2) The map is built on board and the waypoints of the safe path are also calculated on board.
  - OA3: (W=1) The robot transmits live positions and images/video to the control station.
- Set OPI
  - 1 sign for the machine room; X blocked paths; Y indoor obstacles; Z damages.

Required Data:

- Built map with OPI marks and positions as well as the estimated path.
- OPI images and associated positions.
- Timing data.

Benchmarked Functionalities:

- 2D Mapping
- Object recognition

4 Functionality Benchmarking for euRathlon 2015

In Section 3 above, we have defined a general framework for ranking and scoring for Task Benchmarks and applied this framework to each of the tasks. It is not possible to define a single scoring framework for all Functionality Benchmarks because each functionality is specialised and the metrics for different functionalities would be significantly different. So this Section will define the scoring methodologies and metrics separately for each Functionality Benchmark of euRathlon 2015 competition.

The Functionality Benchmarks consist of following four parts:

- Description: This will give a general, high level description of the functionality.
- Input/Output: The Input part will describe what data are needed in order to execute the module implementing the functionality.
- Benchmarking data: This will describe the data needed to perform the evaluation of the functionality, e.g., the sensor data provided by the robots, the ground truth data and/or the result information such as the number of OPIs found.
- Metrics: The Metrics part will describe the methods or algorithms used to process the benchmarking data.

Note: in the current competition plan the Functionality Benchmarking experiments are not designed separately to evaluate individual functionalities. Instead they are implemented in the Task Benchmarks, so the data required for the Functionality Benchmarks will be collected during the task executions. And the Functionality Benchmarking will be the post-processing after the competition.

In the following Sections we will define the Functionality Benchmarks in detail for euRathlon 2015 based on these four aspects.

4.1 2D Mapping

Description:
Starting from a given vehicle position (this can be the initial position of the vehicle using GPS, build a map of the environment. In this context, a map is defined as "any digital representation of the environment suitable for performing other functionalities (e.g., self-localization, path planning, object recognition/detection etc.)". Depending on the specific robot platform under test, mapping requires a more or less extended exploration of the environment.

Input/Output:
The input is the sensor data provided by the devices of the robot system under test. The expected output is the representation of the environment, which is most suitable for the intended use.

Benchmarking data: a set of known physical points in the environment and a map representation provided by the robot.

Metrics:

- **Map Coverage (MC):** This is simply the % of the provided map coverage compared to the expected minimum coverage provided by euRathlon. The % cannot exceed 100%.
- **Metric Accuracy (MA):** we propose the Root-Mean-Square Error (RMSE) between real (ground truth) x, y positions and the robot's estimated x, y positions of the same features. This is measured in mm, and it is a real value metric – the lower the better.

Note: 2D Mapping will not be done for Sea domain due to dynamic environment conditions.

Benchmarking example:

Suppose we have 3 features in the map, and the known (x, y) positions, with reference to a fixed origin or datum point, for these features are, p_1 = (1, 1); p_2 = (2, 4) and p_3 = (4, 4), all in metres. Then the robot, using the same fixed origin, maps the terrain and locates the same three features – perhaps using SLAM or a related approach. The robot's position estimates for these features will be p'. Suppose that p'_1 = (0.9, 1.1), p'_2 = (2.05, 3.8) and p'_3 = (4.1, 3.95).

The Root Mean Square Error is calculated as follows:

The Euclidian distance (squared) between p_i and p'_i, is

$d_i = (x_i - x'_i)^2 + (y_i - y'_i)^2$

Thus for the example above, with 3 points:

RMSE $= 1/3 * SQRT((d_1 * d_1 + d_2 * d_2 + d_3 * d_3)) = 0.158$m

Root Mean Square Error						
	Ground truth		Estimates		Euclidian distance	
p	x	Y	x	y	d_i	SQRT(d_i)
1	1	1	0.9	1.1	0.02	0.14142136
2	2	4	2.05	3.8	0.0425	0.20615528
3	4	4	4.1	3.95	0.0125	0.1118034
				RMSE	0.158m	

From this error, a metric index is calculated as follows:
Index = tanh(MaxError-RMSE) if RMSE<MaxError and 0 otherwise. This function normalises the index between [0 and 1] for all errors less than a Maximum Error allowable. The value of MaxError will be fixed based on competitor's performances.
A map index value in [0, 1] will be computed using the mean of the metric index and the Map Coverage (MC).
For example, assume that the maximum allowable RMSE is 2 meters and the map coverage is 75% for team A, the team would get a mapping index of (0.95+0.75)/2 = 0.85.
Assume now that team B has an overall RMSE of 0.4 with a 100% map coverage, its mapping index would be 0.96.

4.2 Object recognition

Description:

A robot perceives the presence of the specific objects in the environment by making use of its sensors. Objects belong to a certain class or category (for example, workers, debris, vegetables etc.). Each object is a particular instance of the class, for example a worker is an instance of workers class. The Object Recognition here includes the object detection which identifies the presence of the class and the actual object recognition that identifies which class the object belongs to (e.g. a worker, a pipe, a valve or an obstacle). The Object Recognition functionality also needs to do the object localization which gives the location information of the object.

Input/Output:

The input data will be the sensor data from the devices of the robot system under test. For detecting the object class the set of categories the objects belong to and the description of each object (e.g. a simple label of the object) are known to the robot beforehand. The output is a yes/no in terms of the class presence, which is used for further object recognition which the expected output would be the label of the object. The output also includes position and/or orientation information. The imagery of the object also needs to be acquired in euRathlon2015.

Benchmarking data:

Given a list of possible objects the robot should return the label of the class for a specific object. The location information and the imagery should also be returned. The benchmarking can be run during the competitions or by providing a pre-record stream of sensory input.

Metrics:

Inspired from domains of Pattern Recognition and Information Retrieval, several metrics could be used [3]:

- **Accuracy:** Accuracy is measured as the percentage of correctly recognized objects and correctly rejected objects against all the objects that are presented to the robot. This is a real-valued metric, the higher the better. This measurement is not used in euRathlon 2015 because it is difficult to know the number of all objects including various obstacles in the outdoor domain.

- **Precision and Recall:** The Precision is measured as the percentage of correctly recognized objects against all recognized objects which include correct ones and incorrect ones. The Recall (also known as sensitivity) is measured as the percentage of correctly recognized objects against all objects that should have been detected in the scene.

  For example, suppose there are 4 missing workers, 5 pieces of debris and 6 big stones presented in the scene the task is to recognize the missing workers. If 3 objects are recognized but only 2 objects are workers (correctly recognized) and other two recognized objects may be a piece of debris and a stone (incorrectly recognized), the precision will be 2 / 3 = 66.66%, the recall will be 2 / 4 = 50.00%

  For the accuracy, the number of the correctly recognized objects is 2 and the number of correctly rejected objects is (5+6) - 2 (recognized incorrectly as workers) = 9. The accuracy is (2+9)/(4+5+6) = 73.33%

Precision can be seen as a measure of exactness or quality, whereas recall is a measure of completeness or quantity.

- **F-measure:** Both precision and recall are real-valued, though the higher the better, each high one could mislead the interpretations. To consider them together a combined measure, the traditional F-measure or balanced F-score which is the harmonic mean of precision and recall, is used.
  The general form of F measure is

$$F_\beta = \left(1 + \beta^2\right) \cdot \frac{precision \cdot recall}{\beta^2 \cdot precision + recall}$$

Formula 3. General form of F-measure

where β is a non-negative real value.

The traditional F-measure or balanced F-score (F_1 score) is the harmonic mean of precision and recall.

Two other commonly used F measures are the F_2 measure, which weights recall higher than precision, and the $F_{0.5}$ measure, which puts more emphasis on precision than recall.

The F_1-measure will be used. It is a real-valued metric, the higher the better

- **Position Error:** it is used to evaluate the performance of the object localization. The distances between the localized position and the real position could be used. It is a real value; the smaller the better.

- **Time:** the time required to execute the recognition/identification/localization, including the time to process the data input. It is a real value; the smaller the better.

4.3 Object Manipulation

Description:

In some scenarios a robot is required to manipulate objects, e.g. open or close a series of valves to stem pipe leaks, pick up a canister from ground and put it into a barrel for contaminated materials.

Input/Output:

The model data of the object could be provided to the robot. The robot is expected to locate the object and perform the operation, e.g. closing the valve.

Benchmarking data:

The data required to evaluate will be a recording of the scene while the operation is performed. This will be done by the organisation which will position fixed cameras around the manipulation area.

Metrics:

- **Success:** How well the robot performs the required manipulations in range [0,1]. For example, 0 corresponds to the situation where the valve was grabbed but not turned, 1 is when the valve has been fully turned and 0.5 is when the valve is half turned. The same applies to the canister manipulation. This can be measured by a fixed camera view before and after manipulation (provided by the organization). This considers the mistakes the robot makes, e.g. not fully closes a valve, picks up the canister but not put it into the barrel (Note that picking up a wrong object e.g. a stone should be considered as an error of the functionality of Object Recognition rather than Object Manipulation).

  A real value metric in the range [0,1], the higher the better.

- **Time:** The time taken to perform the manipulation. A real-valued metric, the smaller the better.

4.4 Obstacle Avoidance

Description:
Scenarios in Land and Sea domain there will be static obstacles (i.e. debris, stones, holes, vegetation, etc.) and dynamic obstacles (i.e. sea life). A robot needs to avoid the obstacles in order to successfully perform the tasks.

Input/Output:
The model data of the object could be provided to the robot. The robot is expected to pass by the obstacles.

Benchmarking data:
Performed by the judges. If the Navigation data is available, this will also be overlaid on the map of the area and compared with the identified free paths available to the robots.

Metrics:

- **Success:** The number of times a robot hits the obstacles. It is an integer value, the smaller the better.

5 Benchmarking Data Formats

For each task defined in Section 3, the teams are required to provide data that will be used to perform the benchmarking. These data are categorised as Vehicle Navigation Data, Mission Status, Mapping Information and Object Recognition Information and their formats required are described as below.

5.1 Vehicle Navigation Data

The Vehicle Navigation Data should be in kml format and has following requirements:

- The data sampling frequency: 1 Hz, i.e. a data sample every one second.
- Time: UTC time
- Position: Latitude, Longitude (in decimal degrees)
- Heading: (in degrees)
- Depth: Sea Domain only (in meters)
- Altitude: Air/Sea domains (in meters)

5.2 Mission Status Data

This will give the information related to the status of the mission undertaken should be in kml format with the following requirements:

- Subtask undertaken: Text
- Key decision message: Text
- Time: UTC time. Should be a series of Time corresponding to a series of events, e.g. the subtask starts, the subtask ends, start to close a valve, finish closing the valve, etc. The Time can be used as one of the measurements for Functionality Benchmarking (e.g. for Object Manipulation benchmarking).

5.3 Map Information

The map information submitted by the teams should include following information and formats:

- The map file: (KML format – Keyhole Markup Language). KMZ files with a kmz extension
- Abstract Level information: OPIs, Features. This should be integrated in the kmz file
- 2D/3D map in raster or vector format with geo-reference information for high bandwidth data.

5.4 Object Recognition Information

The Object Recognition information should be stored in kml format and include the following:

- Target ID: Text/Number
- Target position (Latitude, Longitude, Depth)
- Target image: image files (png, jpeg etc.)

All information (map, navigation, mission status) should be stored in a single kml file if possible.

All data submitted will be used for the Functionality Benchmarking and as shared data as planned for the euRathlon competition.

6 Bibliography

[1] "RoCKIn robot competition project," 2014. [Online]. Available: http://rockinrobotchallenge.eu/.

[2] "euRathlon 2015 Scenarios," 2015. [Online]. Available: http://www.eurathlon.eu/site/index.php/compete/eurathlon-2015-scenarios/.

[3] "Precision and Recall," [Online]. Available: http://en.wikipedia.org/wiki/F1_score.